




Construção e Interpretação de Diagramas de Extremos e Quartis por Alunos Portugueses do 9.º ano de Escolaridade

Building and Interpreting Boxplots by 9th Grade Portuguese Students

Maria José Carvalho*

 ORCID iD 0000-0003-2045-9612

José António Fernandes**

 ORCID iD 0000-0003-2015-160X

Adelaide Freitas***

 ORCID iD 0000-0002-4685-1615

Resumo

Este estudo debruça-se sobre as dificuldades de alunos portugueses do 9.º ano de escolaridade na construção e interpretação de diagramas de extremos e quartis (DEQ). O estudo tem por base as respostas dadas por 93 alunos do 9.º ano a duas questões sobre DEQ num teste de diagnóstico escrito. Examinando as configurações dos objetos e processos matemáticos que intervêm nas resoluções dos alunos, identificaram-se discrepâncias entre os significados pessoais atingidos pelos alunos e os significados institucionais pretendidos na construção e interpretação de DEQ, constatando-se que: (i) na construção de DEQ observam-se confusões na forma de apresentar um DEQ ao incorporar o seu suporte numérico no próprio diagrama, detectando-se dificuldades na precisão a tomar para a escala do suporte quando o domínio de variação dos dados é maior; e (ii) na interpretação de DEQ são manifestas dificuldades em analisar a forma distribucional dos dados através de um DEQ, observando-se a tendência para o aluno responder com base na perceção intuitiva da simetria e dispersão da distribuição, sem recorrer a medidas estatísticas adequadas na sua argumentação. Observou-se, ainda, que a construção do DEQ é mais fácil do que a sua interpretação. Estes resultados corroboram a ideia de que a análise de dados agregados, usando DEQ, não é uma tarefa fácil de compreender por parte de alunos deste nível de escolaridade, sugerindo-se uma maior articulação entre atividades de construção de DEQ e atividades de interpretação.

Palavras-chave: Conflitos Semióticos. Diagrama de Extremos e Quartis. Alunos do 9.º ano.

Abstract

This study deals with the difficulties of 9th grade Portuguese students in the building and interpreting of boxplots

* Doutor em Educação, na especialidade de Didática e Desenvolvimento Curricular pela Universidade de Aveiro (UA). Professora e Formadora no Agrupamento de Escolas D. Pedro I (AE D. Pedro I) Vila Nova de Gaia, Portugal. Endereço para correspondência: Rua Nova do Fojo, 4400-232 Vila Nova de Gaia, Portugal. E-mail: maria.carvalho@ua.pt.

** Doutor em Educação pela Universidade do Minho (UM). Professor Associado no Instituto de Educação da Universidade do Minho (UM), Braga, Portugal. Endereço para correspondência: Campus de Gualtar, 4710-057, Braga, Portugal. E-mail: jfernandes@ie.uminho.pt.

*** Doutor em Matemática pela Universidade de Aveiro (UA). Professora auxiliar no Departamento de Matemática (DMat) e membro do Centro de Investigação e Desenvolvimento em Matemática e Aplicações (CIDMA) da Universidade de Aveiro, Aveiro, Portugal. Endereço para correspondência: Campus Universitário de Santiago, 3810-193 Aveiro, Portugal. E-mail: adelaide@ua.pt.

(DEQ). The study is based on the responses given by 93 9th grade students to 2 questions on the DEQ in a written diagnostic test. By examining the configurations of the objects and mathematical processes that intervene in the students' resolutions, we have identified discrepancies between the personal meanings reached by the students and the institutional meanings intended in the construction and interpretation of the DEQ, noting that: (i) in the construction of the DEQ we observed confusions in the form of presenting a DEQ by incorporating its numeric support in the diagram itself, and difficulties in the accuracy to be considered for the support scale of the support when the data's variation domain is greater; and (ii) in the interpretation of DEQ, there are clear difficulties in analyzing the distributional form of the data through a DEQ, observing the tendency for the student to respond based on the intuitive perception of the symmetry and dispersion of the distribution, without resorting to statistical measures in their justifications. We also observed that the construction of the DEQ is easier than its interpretation. These results corroborate the idea that the analysis of aggregate data, using DEQ, is not an easy task to understand by students of this level of education, so we suggest a greater articulation between DEQ construction activities and interpretation activities.

Keywords: Semiotic Conflicts. Boxplot. 9th Grade Students.

1 Introdução

Atualmente, a Estatística é considerada uma área fundamental em várias dimensões da vida das pessoas, tais como a vida quotidiana, o emprego e a participação social. Ao mesmo tempo, também a sociedade contemporânea requer cidadãos habilitados com maiores capacidades de leitura, organização e interpretação de informação. Esta necessidade exige o desenvolvimento da literacia estatística, que segundo Rumsey (2002) engloba dois tipos de resultados de aprendizagem dos alunos: “ser capaz de funcionar como um membro educado da sociedade nesta idade da informação e ter uma compreensão significativa de termos, ideias e técnicas estatísticas” (p. 3).

Ora, a valorização da literacia estatística tem-se repercutido na maior importância dada ao ensino da Estatística nas escolas nos últimos 30 anos. No caso do ensino básico, em Portugal, que decorre entre o 1.º e o 9.º ano, o tema de Estatística, que se insere no domínio Organização e Tratamento de Dados (OTD), faz atualmente parte dos programas portugueses de todos os anos escolares da disciplina de Matemática (MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA, 2013). Segundo os programas, o ensino da Estatística visa desenvolver nos alunos capacidades de interpretação, de avaliação crítica e de comunicação de informação estatística (GAL, 2002; OLIVEIRA; HENRIQUES, 2014).

Para Lopes (2004), o ensino da Estatística poderá contribuir para o desenvolvimento das capacidades crítica e de autonomia, capacidades que estão indissociavelmente relacionadas com a construção e interpretação de gráficos estatísticos e, consequentemente, com a literacia estatística (MORAIS; FERNANDES, 2011).

Assim, refletindo a comunicação social, enquanto meio no qual a literacia estatística está muito presente (ARTEAGA; BATANERO; CAÑADAS; CONTRERAS, 2011; ESPINEL;

GONZÁLEZ; BRUNO; PINTO, 2009), o programa de Matemática incentiva a construção e interpretação de tabelas e gráficos estatísticos desde o primeiro ano de escolaridade (MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA, 2013), valorizando-se assim a representação gráfica no ensino e na aprendizagem da Estatística.

Potenciar ao máximo um gráfico, segundo Curcio (1989), é conseguir interpretá-lo e extrair o máximo de informação possível sobre os dados nele representados. Nos últimos anos tem-se observado um aumento do número de estudos que investigam a compreensão de gráficos por parte dos alunos. Alguns estudos apontam dificuldades dos alunos em representarem e relacionarem quantidades através de gráficos (PAIS; SARAIVA, 2011), outros insistem no porquê dessas dificuldades (PONTE; FONSECA, 2001; SERRANO; ORTIZ; RODRIGUEZ, 2009) e, ainda, alguns focam-se na aptidão dos alunos para lerem e identificarem fatores que influenciam a compreensão de gráficos (FRIEL; CURCIO; BRIGHT, 2001).

Este estudo está centrado num tipo particular de representação estatística: o diagrama de extremos quartis (DEQ), que alguns autores (BAKKER; BIEHLER; KONOLD, 2004; MORAIS; FERNANDES, 2011; MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA, 2013) consideram como sendo um tipo de gráfico de extrema relevância, já que sumaria, numa só representação, a centralização e a dispersão dos dados. Mais concretamente, neste estudo analisam-se as respostas dadas por alunos portugueses do 9.º ano de escolaridade a questões envolvendo a construção e interpretação de DEQ.

Seguidamente, nas próximas seções, apresenta-se o enquadramento teórico do estudo, a metodologia seguida na sua implementação, os resultados obtidos e, por último, reflete-se sobre as principais conclusões do estudo e discutem-se pistas para melhorar a aprendizagem dos alunos.

2 Enquadramento teórico

2.1 Marco teórico

O modelo teórico do Enfoque Ontossemiótico (EOS) do conhecimento e do ensino da Matemática, que Godino e colaboradores têm vindo a desenvolver (e.g., GODINO, 2011; GODINO; BATANERO, 1994; GODINO; BATANERO; FONT, 2007; GODINO et al., 2017), envolve cinco noções teóricas básicas, as quais se revelam úteis na análise dos processos de ensino e aprendizagem, que são: *sistema de práticas; configuração de objetos e processos matemáticos, emergentes e intervenientes nas práticas matemáticas; configuração didática;*

dimensão normativa; e noção de idoneidade didática (GODINO et al., 2017). No presente estudo são especialmente relevantes as noções de *sistemas de práticas e configuração de objetos e processos matemáticos*.

A noção de *sistema de práticas* implica uma visão antropológica e pragmática da Matemática, permitindo introduzir as noções de significado institucional e pessoal dos objetos matemáticos (GODINO; BATANERO, 1994). De acordo com o EOS, a aprendizagem envolve a apropriação, pelos alunos, dos significados institucionais através da sua participação em práticas sociais geradas na sala de aula, envolvendo o ajustamento progressivo entre os significados pessoais iniciais do aluno e os significados institucionais planeados. Neste modelo, os significados são entendidos em termos de práticas operacionais e discursivas, envolvendo também o reconhecimento e a interação dos objetos implicados nessas práticas.

Na noção de *configuração de objetos e processos matemáticos* adota-se uma perspetiva interacionista de objeto e pragmática de significado. A configuração ontossemiótica, envolvendo práticas, objetos e processos, dá resposta à necessidade de identificar os objetos e processos usados nas práticas matemáticas de resolução de situações-problema. Estas configurações constituem uma ontologia de objetos matemáticos primários, os quais possibilitam uma análise aprofundada e ampla das práticas mobilizadas nos processos de resolução de problemas. Para tal, Godino, Batanero e Fonte (2007), na análise da atividade matemática, consideram importante considerar seis tipos de objetos primários: *situações-problema* – aplicações extramatemáticas, exercícios, problemas, ações que induzem uma atividade matemática; *linguagens* – termos, expressões, notações, gráficos que se utilizam para representar os dados de um problema, as operações que se efetuam, os objetos matemáticos que se usam e a solução encontrada em seus diversos registos; *conceitos/definições* – formulações introduzidas mediante definições e descrições, as quais constituem práticas realizadas pelos alunos para resolver um problema matemático com uso implícito ou explícito de objetos matemáticos, que o aluno tem de lembrar e aplicar; *propriedades/proposições* – enunciados sobre relações ou propriedades dos conceitos que se utilizam para resolver problemas matemáticos; *procedimentos* – algoritmos, operações, técnicas de cálculo que os alunos aplicam para a resolução do problema; e *argumentos* – enunciados usados nas respostas para validar ou explicar as proposições e procedimentos ou a solução dos problemas, que podem ser dedutivos, indutivos, formais ou informais. As configurações ontossemióticas, do ponto de vista institucional, são designadas por configurações epistémicas, e de um ponto de vista pessoal, são designadas por configurações cognitivas.

Estes seis objetos primários relacionam-se originando configurações de objetos e

diferentes processos matemáticos envolvendo os objetos, descritos em facetas duais: *pessoal-institucional* – consoante o significado emerge da prática pessoal ou de um grupo de pessoas que partilham o mesmo tipo de problemas, ou seja, uma instituição; *ostensivo-não ostensivo* – conforme o objeto é público, pode ser mostrado a outro ou materializado de alguma forma (associado a notações, símbolos, gráficos), ou é pensado, imaginado por um sujeito, ou está implícito no discurso matemático; *extensivo-intensivo* (particular-geral) – utiliza-se na generalização de resultados numa atividade matemática, permitindo centrar a atenção na dialética entre o particular e o geral; *unitário-sistémico* – em certas circunstâncias, os objetos matemáticos participam como entidades unitárias, enquanto noutras devem ser tomados como sistemas para que se possa proceder ao seu estudo; *expressão-conteúdo* (antecedente e consequente de qualquer função semiótica) – a atividade matemática, os processos de construção e uso dos objetos matemáticos caracterizam-se por serem fundamentalmente relacionais, colocando os distintos objetos em relação uns com os outros. Esta relação estabelece-se por meio de funções semióticas, entendidas como uma relação entre um antecedente (expressão, designação ou nome) e um consequente (conteúdo, designado ou ente matemático) instituído por um sujeito (pessoa ou instituição) de acordo com determinado critério ou código de correspondência.

Em relação às funções semióticas, refira-se que elas são entendidas como relações entre conjuntos, envolvendo três componentes: *expressão*, que constitui o objeto inicial ou significante; *conteúdo*, que é o objeto final ou significado; e *regra de correspondência*, que é o código interpretativo que regula a relação entre a expressão e o conteúdo. Esta dualidade expressão-conteúdo possibilita análises semióticas *a priori* (etapa prévia de análise didático-matemática) e *a posteriori* (análise das produções escritas dos alunos numa tarefa matemática).

A discordância entre os significados atribuídos a uma expressão por dois sujeitos (pessoas ou instituições) em interação comunicativa, designado por conflito semiótico, é útil para a análise da produção e comunicação matemática. Se a disparidade se produz entre significados institucionais, falamos de conflitos semióticos do tipo epistémico; se ela se produz entre práticas que formam o significado pessoal de um mesmo sujeito, expressa conflitos semióticos do tipo cognitivo (GODINO; BATANERO; FONT, 2007).

Neste trabalho, a investigação envolve a identificação de dificuldades na compreensão dos DEQ, na sua vertente de construção e interpretação, tendo por base a produção escrita de respostas dos alunos (significado pessoal) a questões sobre DEQ, a análise dos significados dos conteúdos das funções semióticas demonstrados nessas respostas, comparando-os com os significados institucionais de referência e a cadeia de procedimentos prevista no processo

correto de resolução das questões. A verificação de discrepâncias entre o significado institucional e o significado pessoal (GODINO; BATANERO, 1994) levará a identificar conflitos semióticos.

2.2 Antecedentes

O DEQ é um tipo de representação estatística de dados que surgiu nos anos 70-80 do século passado, quando dos desenvolvimentos da Estatística verificados no âmbito da Análise Exploratória de Dados (TUKEY, 1977). A Análise Exploratória de Dados constitui uma primeira abordagem dos dados, com vista a descrever, gráfica e numericamente, uma coleção de dados e, desse modo, identificar regularidades e tendências nessa coleção de dados. Com essa análise preliminar dos dados espera-se então potenciar o estabelecimento de conjecturas sobre a população de onde foram extraídos os dados, as quais poderão ser avaliadas ou testadas por meio de outros recursos estatísticos (métodos inferenciais).

A construção do DEQ implica o conhecimento de cinco medidas estatísticas da coleção de dados: o valor mínimo (mín), o 1.º quartil (Q_1); o 2.º quartil (Q_2 ou mediana); o 3.º quartil (Q_3) e o valor máximo (máx). Seguidamente, representando esses valores na reta numérica (suporte numérico do DEQ), construindo o retângulo delimitado pelo 1.º e 3.º quartil e assinalando os valores mín e máx, obtém-se a representação pretendida (Figura 1).

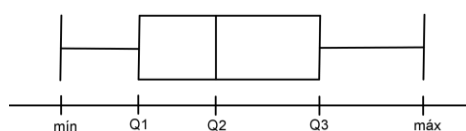


Figura 1 – Diagrama de extremos e quartis.
Fonte: Elaboração dos autores (2018).

Conforme é evidenciado na literatura (e.g., BAKKER; BIEHLER; KONOLD, 2004; EDWARDS; ÖZGÜN-KOCA; BARR, 2017), construir um DEQ conhecidos os cinco valores referidos não constitui uma dificuldade considerável para os alunos. Antes, as dificuldades dos alunos podem resultar da necessidade de determinar os três quartis, até porque a determinação da mediana tem-se revelado mais difícil para os alunos do que a determinação da média e da moda (BOAVENTURA; FERNANDES, 2004; FERNANDES; BARROS, 2005).

Tendo em conta a definição dos quartis, conclui-se que este tipo de representação tem uma propriedade muito particular: entre quaisquer dois elementos consecutivos, dos cinco representados num DEQ, tem-se sempre cerca de 25% de dados. Ora, esta propriedade pode ser usada para avaliar a concentração/dispersão dos dados (agregados) em diferentes intervalos de

extremos envolvendo os valores referenciados no DEQ e, em consequência, permite extrapolar sobre a simetria/assimetria da distribuição empírica correspondente aos dados que representa.

Embora aparentemente simples, esta propriedade dos DEQ não é facilmente compreendida pelos alunos, que tendem a considerar que os intervalos (entre elementos consecutivos referenciados) com maiores amplitudes contêm também um maior número de dados. Segundo Edwards Özgün-Koca e Barr (2017), este tipo de representação de dados é contraintuitivo, no sentido de que as aprendizagens anteriores sobre outros tipos de representações gráficas não contribuem para as novas aprendizagens implicadas nesta representação.

No DEQ, especificamente, Bakker Biehler e Konold (2004) salientam que a extensão dos intervalos está inversamente relacionada com a densidade dos dados, isto é, as amplitudes dos intervalos serão tanto menores quanto mais próximos estiverem os dados uns dos outros. Deste modo, as frequências dos dados não são salientadas neste tipo de representação, tal como acontece nos diagramas circulares que os alunos aprenderam antes (6.º ano de escolaridade), nos quais setores circulares maiores correspondem a maiores frequências. Diferentemente, nos DEQ podem existir intervalos com amplitudes diferentes para a mesma quantidade de dados e quantidades diferentes para amplitudes iguais. Por exemplo, no DEQ da Figura 2 existem sempre três observações entre valores consecutivos de referência (mín, Q_1 , Q_2 , Q_3 , máx); mais ainda, a amplitude interquartil ($Q_3 - Q_1$) é igual à amplitude entre o 3.º quartil e o valor máximo, mas no primeiro estão incluídos seis observações e no outro apenas três.

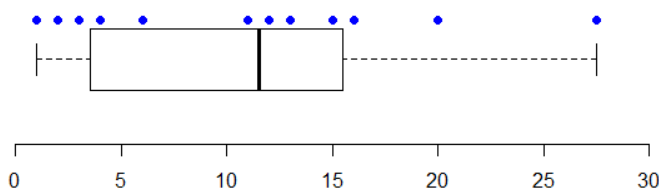


Figura 2 – DEQ relativo à coleção de dados: 1; 2; 3; 4; 7; 11; 12; 14; 15; 16; 20; 27,5.
Fonte: Elaboração dos autores (2018).

Por outro lado, no DEQ perde-se a informação dos dados *per si*. Com base somente na informação agregada dos dados, Edwards Özgün-Koca e Barr (2017) constataram que a interpretação de um DEQ se torna mais vulnerável aos alunos. Num estudo envolvendo 259 alunos do 6.º e 8.º anos de escolaridade, aqueles autores verificaram ser mais provável encontrar respostas mais desenvolvidas sobre percentagens de observações acima ou abaixo de medidas de referência nos DEQ quando os alunos têm também acesso aos dados originais, pois estes podem ser usados para obterem a sua resposta. Isso, contudo, não implica uma resposta correta por dificuldades inerentes, muitas vezes, ao conceito de percentagem segundo aqueles autores.

Tendo também a informação dos dados não agregados, importa salientar que a existência de observações repetidas na coleção de dados pode igualmente condicionar o sucesso do aluno na interpretação de um DEQ. Na realidade, à repetição de observações, embora comum em problemas propostos em manuais escolares, não tem sido dada a devida atenção nos próprios manuais (FREITAS; CRUZ; SILVA, 2017), não se alertando para o pormenor de poderem existir observações iguais, em particular, iguais aos valores de referência (mín, Q_1 , Q_2 , Q_3 , máx) de um DEQ e o que tal implica na interpretação da mediana e dos quartis, como é discutido por Freitas, Cruz e Silva (2017).

Na Figura 3 são apresentados três DEQ iguais (com mín=1; $Q_1 = 1,2$; $Q_2 = 2$; $Q_3 = 3$; máx=5), correspondentes a três coleções distintas de dados, sendo que duas coleções contêm observações repetidas e iguais aos quartis. Para os três DEQ poder-se-á dizer que entre o Q_1 e o Q_2 estão cerca de 25% dos dados, mas nos dois primeiros diremos que pelo menos 25% dos dados da coleção são valores entre 1 e 1,2, sendo que apenas no último DEQ estão exatamente 25% dos dados.

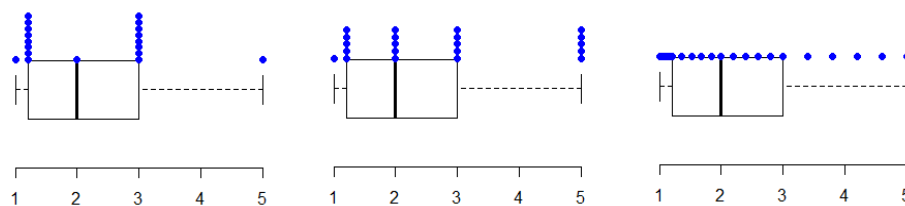


Figura 3 – DEQ iguais para três coleções distintas, com $n = 19$ observações.
 Fonte: Elaboração dos autores (2018).

Curcio (1989) estabeleceu uma taxonomia envolvendo três níveis de leitura e interpretação de gráficos, a qual permite clarificar as maiores/menores dificuldades dos alunos. No primeiro nível, *ler os dados*, espera-se que o aluno identifique dados apresentados explicitamente no gráfico, o que se realiza através da leitura de fatos que nele estão representados. Neste nível não há interpretação dos gráficos e pretende-se que o aluno compreenda a escala e as unidades de medida. No caso dos DEQ, indicar as cinco medidas estatísticas de referência corresponde a este primeiro nível. No segundo nível, *ler entre os dados*, espera-se que o aluno interprete e organize a informação fornecida pelos dados. Neste nível, o aluno deve combinar e integrar a informação e identificar relações matemáticas através de algum conhecimento prévio sobre o assunto tratado no gráfico. Como refere Curcio (1989), este é o nível mais comum na resolução de tarefas que envolvem a compreensão dos gráficos, esperando-se que o aluno identifique tendências no gráfico e relacione dados. No caso dos DEQ, questões focadas na estrutura dos DEQ como, por exemplo, identificar a percentagem de dados acima de um valor dado (por exemplo, o primeiro, segundo ou terceiro quartil) ou comentar a

dispersão dos dados tomando a amplitude total e a amplitude interquartil correspondem a este segundo nível de Curcio.

Finalmente, no terceiro nível, *ler para além dos dados*, pressupõe-se que o aluno, ao ler a informação do gráfico, infira a informação total e tenha um conhecimento prévio aprofundado sobre o assunto referente aos dados do gráfico. Neste nível, o aluno deve conseguir responder a questões cujas respostas requerem o uso de informação implícita no gráfico, extrapolando, predizendo ou fazendo inferências. Ou seja, como refere Curcio (1989), pretende-se que o aluno se projete no futuro e coloque questões sobre os dados. No caso dos DEQ, correspondem ao terceiro nível questões sobre, por exemplo, a previsão do valor de um dado sabendo que ele excede o terceiro quartil, ou, ainda, sobre a simetria/assimetria da distribuição, supondo que a distribuição (desconhecida) entre valores de referência do DEQ seja uniforme (como é suposto no DEQ à direita na Figura 3).

Em geral, segundo Friel, Curcio e Bright (2001), os alunos não têm dificuldades no primeiro nível, ler os dados, e revelam dificuldades nos dois níveis seguintes: ler entre os dados e ler além dos dados. No estudo com alunos do 6.º e 8.º ano, antes referido, Edwards Özgün-Koca e Barr (2017) concluíram que os alunos foram capazes de ler no diagrama qualquer dos cinco valores subjacentes à sua construção, o que corresponde ao primeiro nível de Curcio (1989), mas exibiram muitas dificuldades no segundo nível de Curcio (1989) ao analisar a estrutura de DEQ, nomeadamente na identificação da percentagem de dados acima do primeiro quartil.

As grandes dificuldades que têm sido reveladas pelos alunos, sobretudo na interpretação dos DEQ, levaram Bakker Biehler e Konold (2004) a questionar o seu ensino a alunos muito jovens, argumentando que: 1) os DEQ não permitem, em geral, perceber casos individuais; 2) os DEQ operam de modo diferente das representações gráficas que os alunos encontraram antes; 3) a mediana não é uma estatística tão intuitiva para os alunos como foi pressuposto; e 4) os quartis dividem os dados em grupos de uma maneira que poucos alunos (ou mesmo professores) realmente compreendem.

Em Portugal, de acordo com o Programa e Metas Curriculares (MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA, 2013), no 9.º ano de escolaridade, que corresponde ao final do ensino básico, espera-se que o aluno saiba identificar, representar, tratar e analisar conjuntos (simples) de dados de variáveis discretas e resolver problemas envolvendo esses conhecimentos. O tópico DEQ faz parte do programa de Matemática do 8.º ano de escolaridade, constatando-se que esse tópico é, em geral, abordado na última semana de aulas. Tal prática, provavelmente, decorre deste tópico ser o último do programa e, em consonância, de os manuais

escolares mais usados nas escolas também situarem esse tema no último capítulo do manual (FREITAS et al., 2018).

Face à revisão de literatura reportada, o presente trabalho persegue o propósito de investigar o desempenho de alunos na construção e interpretação de DEQ, no ano escolar seguinte à leção dos DEQ e antes da leção prévia de qualquer outro conteúdo de Estatística. Em termos inovadores, apresenta-se um estudo empírico no qual se averigua:

- a influência da forma de apresentação dos dados na construção do respectivo DEQ, avaliada através da comparação de dificuldades dos alunos perante dados fornecidos extensivamente, por meio da descrição de todos os valores numéricos que constituem a coleção de dados e, graficamente, por meio de um gráfico de barras. Neste último caso, poder-se-á dizer que se averigua o primeiro nível de Curcio (1989), no sentido de que se verifica se o aluno identifica os dados a apresentar explicitamente na representação gráfica pedida; e,
- no segundo e terceiro níveis de Curcio (1989), a capacidade de analisar a dispersão/concentração e a simetria da distribuição subjacente aos dados a partir da interpretação do DEQ.

3 Metodologia

Para estudar o desempenho dos alunos na construção e interpretação de DEQ desenvolveu-se um estudo empírico, de natureza, fundamentalmente, quantitativa e de tipo descritivo, na medida em que se avaliou uma realidade pré-existente (GALL; GALL; BORG, 2003), com vista a identificar as dificuldades mais frequentes e os objetos matemáticos primários mais frequentemente associados às dificuldades de alunos do 9.º ano de escolaridade em duas situações distintas: construção e interpretação de DEQ. Na construção de DEQ, pretendeu-se assinalar conceitos/definições e procedimentos usados e comparar as dificuldades quando a situação dos dados fornecidos corresponde a uma coleção não organizada de valores numéricos *versus* dados já organizados num gráfico de barras. Na interpretação de DEQ, perante um diagrama dado, pretendeu-se analisar propriedades e argumentos usados e investigar, comparativamente, a capacidade de análise de DEQ quando, ao ser questionado sobre o comportamento da distribuição subjacente aos dados, não é especificado o tipo de análise a realizar *versus* quando se explicita discutir a dispersão e simetria da distribuição.

Para a implementação do estudo, foram selecionados alunos de duas escolas do distrito do Porto, numa base de amostragem por conveniência. A todas as turmas do 9.º ano das duas escolas, num total de 191 alunos, foi aplicado um mesmo teste de diagnóstico numa aula de

Matemática de 50 minutos e antes do início da leção do tópico Organização e Tratamento de Dados (OTD), com o propósito de obter informação sobre os conhecimentos dos alunos nos conteúdos de Estatística lecionados até ao 8.º ano.

A prova constava de quatro questões sobre as medidas de localização e de dispersão e os DEQ. As duas questões abertas sobre DEQ (aqui designadas por questão 1 e questão 2, ver Figura 4) são o alvo do presente estudo. Ambas incorporam uma alínea para construir um DEQ (no item 1a a partir de dados visualizados num gráfico de barras e no item 2a a partir de dados não organizados) e as outras alíneas focam-se na interpretação do DEQ (no item 1b sem sugerir tópicos a mencionar na interpretação e nos itens 2b e 2c solicitando a interpretação em termos da dispersão e da simetria da distribuição dos dados). Estas questões estão de acordo com os descritores das Metas Curriculares (MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA, 2013) no domínio OTD do 8.º ano.

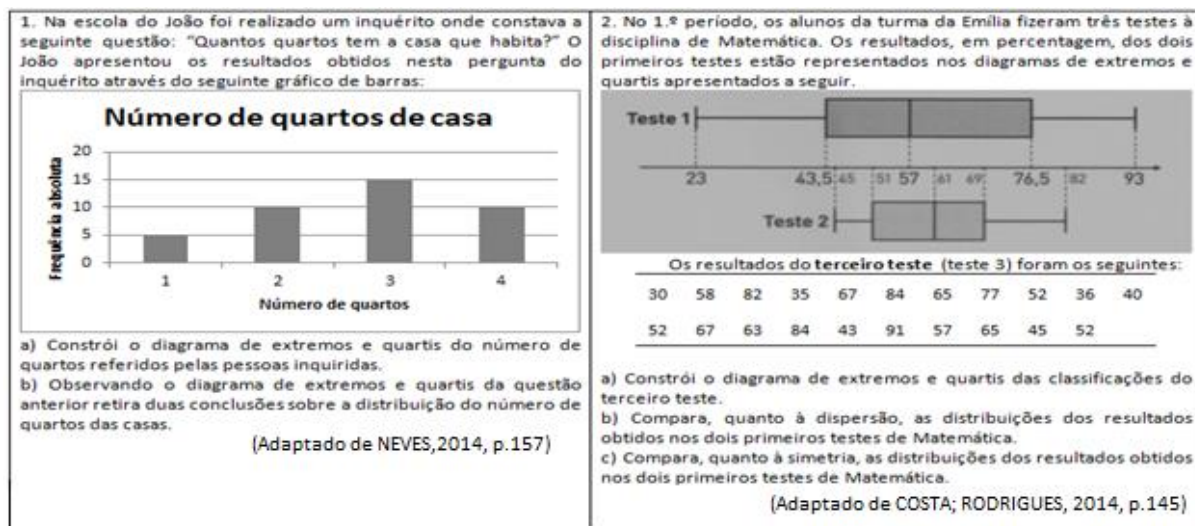


Figura 4 – Enunciados das questões 1 e 2 propostas aos alunos.
Fonte: Elaboração dos autores (2018).

Sendo o tema DEQ lecionado no 8.º ano e os participantes alunos do 9.º ano, tal significa que se avaliou a apropriação dos conhecimentos e as dificuldades demonstradas pelos alunos depois de decorrido cerca de um ano após a sua aprendizagem. Por outro lado, constatou-se, na realização do teste, que vários alunos desconheciam, ou não se lembravam de ter ouvido falar do termo DEQ, donde para este estudo foram eliminadas todas as provas em que não houvesse qualquer registo de respostas, simultaneamente, nas duas questões, 1 e 2, observando-se um total de 98 provas nessas condições. Assim, neste estudo analisaram-se as respostas dadas, às referidas questões, pelos restantes, num total de 93 alunos, com idades compreendidas entre os 13 e os 15 anos e média de idades de 14,3 anos. Destes alunos, 4 usufruíam de apoio educativo

por serem alunos com Necessidades Educativas Especiais e 2 alunos encontravam-se a frequentar o 9.º ano pela segunda vez.

Na perspetiva do EOS, todas as respostas foram alvo de uma análise documental com vista a identificar os objetos e processos usados nas práticas operacionais e discursivas demonstradas nas resoluções. A configuração dos objetos e processos matemáticos permitiram elaborar uma descrição semiótica dos objetos e significados das práticas matemáticas, o que possibilitou o acesso a explicações de eventuais discordâncias ocorridas entre as dimensões pessoal e institucional. Mais ainda, tendo em conta essa análise, cada resposta foi classificada em uma das duas categorias: *resposta correta* e *resposta não correta*, considerando-se ainda a categoria *não resposta*. Finalmente, foram contabilizadas as respostas em cada categoria nas duas questões e construídas tabelas de contingência para pares de alíneas relativas ao mesmo tópico (construção ou interpretação), tendo em vista confrontar as duas situações de cada tópico (item 1a *versus* item 2a, item 1b *versus* item 2b e item 1b *versus* item 2c).

4 Construção e interpretação de diagramas de extremos e quartis

Nesta secção analisam-se as respostas dos alunos à tarefa proposta; em primeiro lugar, aos itens 1a e 2a relativos à construção de DEQ e, seguidamente, aos itens 1b, 2b e 2c relativos à interpretação de DEQ.

4.1 Construção de diagramas de extremos e quartis

Os itens 1a e 2a têm ambos o propósito de investigar as dificuldades dos alunos na construção de DEQ confrontando a situação dos dados serem apresentados num gráfico de barras com os dados serem fornecidos numa coleção não organizada de valores numéricos. Na Tabela 1 apresenta-se a distribuição dos 93 alunos pelo tipo de resposta dada (correta, não correta) e não resposta.

Tabela 1 – Distribuição (%) do tipo de resposta dos alunos nos itens 1a *versus* 2a

Item 1a	Item 2a			Total
	Correta	Não correta	Não resposta	
Correta	6 (6,5)	8 (8,6)	5 (5,4)	19 (20,4)
Não correta	6 (6,5)	53 (57,0)	13 (14,0)	72 (77,4)
Não resposta	2 (2,2)	—	—	2 (2,2)
Total	14 (15,1)	61 (65,6)	18 (19,4)	93 (100)

Fonte: Elaboração dos autores (2018).

Pela Tabela 1 verifica-se que os alunos revelaram dificuldades na construção do DEQ, pois poucos alunos apresentaram respostas corretas (20,4% no item 1a, 15,1% no item 2a e 6,5% simultaneamente nos dois itens). Mais ainda, o fato do item 1a ter recebido uma maior percentagem de respostas corretas (20,4%) e uma menor percentagem de não respostas (2,2%), comparativamente com o item 2a (15,1% e 19,4%, respectivamente), leva a concluir que construir um DEQ quando os dados estão organizados num gráfico de barras constituiu uma tarefa em que os alunos foram mais bem-sucedidos do que no caso dos dados não organizados.

Procurando enfatizar os objetos matemáticos primários que determinam os erros e dificuldades dos alunos na construção do DEQ, foram analisadas as respostas não corretas, considerando-se os três pontos-chave na construção de um DEQ: determinação das cinco medidas estatísticas de referência do DEQ (conceitos/definições e procedimentos); reconhecimento da posição dos valores dessas medidas no suporte numérico do DEQ (procedimento); e saber a forma do diagrama a desenhar sobre o suporte numérico (conceito/definição).

No que concerne ao item 1a, verificou-se que dos 72 alunos com resposta não correta, a maioria deles (61) tiveram dificuldades em determinar o 3.º quartil, sendo que desses, vários indicaram ainda incorretamente o posicionamento da mediana (exemplo na Figura 5) e outros indicaram incorretamente o posicionamento do máximo e do mínimo (exemplo na Figura 6). Constatou-se também que, entre os 91 alunos que responderam ao item 1a, embora o mais comum (27 alunos) tenha sido posicionar corretamente o suporte numérico no DEQ, cerca de um quarto (23 alunos) não desenharam o suporte numérico do DEQ, colocando os valores de referência do DEQ no próprio diagrama (exemplos nas Figuras 5 e 7) e 13 alunos desenharam erradamente o suporte numérico sobrepondo-o ao próprio diagrama (exemplos nas Figuras 6 e 8).

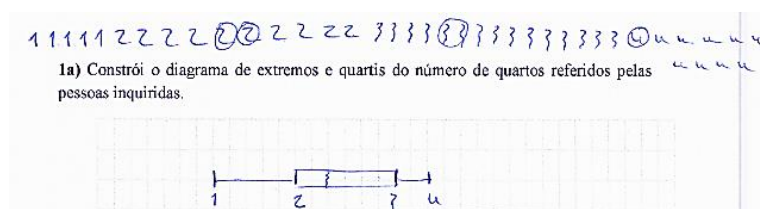


Figura 5 – Resposta do aluno A23 ao item 1a
Fonte: Aluno A23 (2018).

Na resolução apresentada na Figura 5 o aluno revela conhecer a forma do DEQ e saber o conceito/definição e o procedimento de cálculo de Q_1 e Q_2 , contudo erra no valor a atribuir ao Q_3 , dado pelo valor central entre dois valores distintos (contrariamente ao que acontece para

obter Q_1 e Q_2), e ao posicionar incorretamente a mediana quando desenha o seu DEQ. No DEQ reconhece-se falta de rigor na escala usada e não desenha o suporte numérico ao diagrama.

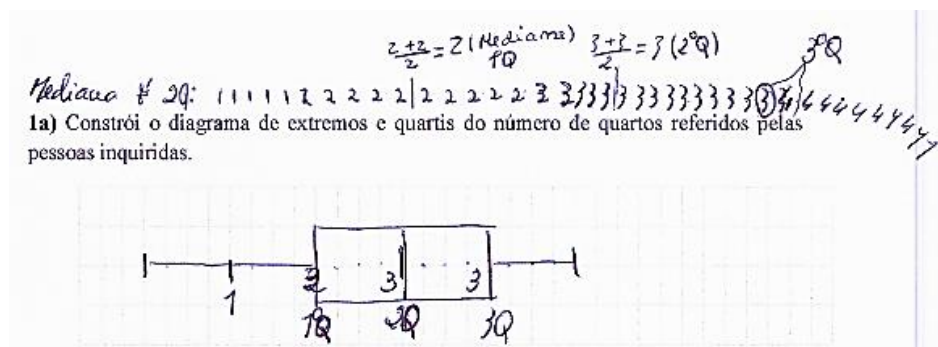


Figura 6 – Resposta do aluno A34 ao item 1a
Fonte: Aluno A34 (2018).

Já na resposta apresentada na Figura 6, o aluno reconhece o algoritmo de cálculo da mediana e dos quartis, mas erra no valor a atribuir a Q_3 , demonstrando alguma confusão nas entidades mediana, primeiro quartil e segundo quartil (refere “mediana” para Q_1 e Q_2), assim como no suporte numérico do DEQ (toma o mesmo número “3” em dois pontos distintos da reta). Mais ainda, posiciona incorretamente ambos os extremos (valor mínimo e valor máximo), não desenhando corretamente o DEQ ao sobrepor os segmentos de reta que unem os quartis aos extremos com a reta de suporte numérico do diagrama, colocando-os, segmentos e suporte, ao mesmo nível horizontal, o que indicia alguma confusão desses objetos no DEQ e torna difícil reconhecer os extremos do DEQ.

No item 2a verificou-se que, dos 61 alunos com respostas não corretas, as dificuldades mais comuns estão no posicionamento do extremo superior no DEQ, com 12 alunos a atribuir o valor máximo possível da variável em estudo, mesmo não tendo sido observado (exemplo na Figura 7), e na colocação do suporte numérico do DEQ, com 18 alunos a incorporar o suporte numérico de apoio à construção do DEQ no próprio diagrama e destacando o valor 90 como valor máximo (exemplo na Figura 8).

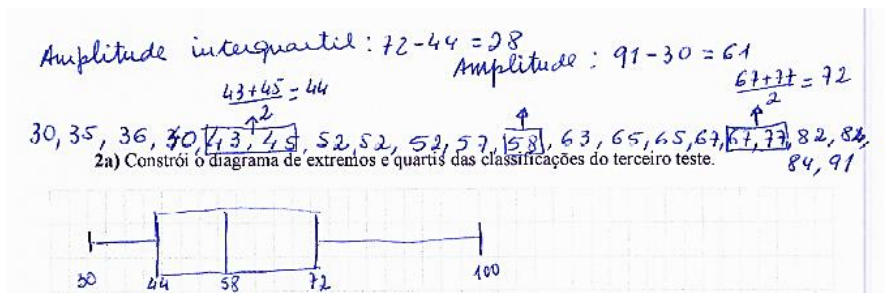


Figura 7 – Resposta do aluno A19 ao item 2a
Fonte: Aluno A19 (2018).

A resposta reproduzida na Figura 7 mostra que o aluno conhece a forma do DEQ, sabe o conceito/definição e o procedimento de cálculo da mediana e dos quartis Q_1 e Q_3 , identifica corretamente os extremos calculando a amplitude total ($\text{máx} - \text{mín} = 91 - 30$) mas, quando constrói o DEQ atribui erradamente o valor 100 ao máximo. Este erro parece revelar confusão entre o valor máximo das classificações efetivamente obtidas no teste com o valor máximo possível da variável (100%). Nota-se também que o aluno não acrescenta o suporte numérico escrevendo, no próprio diagrama, os seus valores de referência.

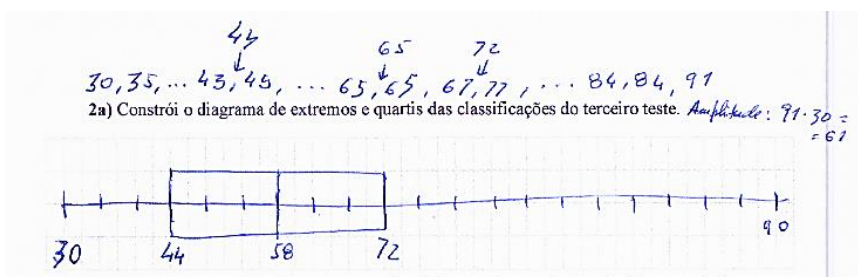


Figura 8 – Resposta do aluno A46 ao item 2a
Fonte: Aluno A46 (2018).

A resposta apresentada na Figura 8, à semelhança da resposta apresentada na Figura 7, não evidencia conflitos a nível dos conceitos/definições e procedimentos nas cinco medidas de referência do DEQ. Porém, e à semelhança da resposta na Figura 6, revela confusão no desenho do DEQ em si, ao sobrepor, no próprio diagrama, o seu suporte numérico (apresentado com uma escala um tanto imprecisa). O aluno, quando efetua a construção do diagrama, destaca o valor “90” e prolonga o traço horizontal, à direita na figura, podendo querer denunciar o número 91 como extremo superior; porém, não o salienta como é requerido no DEQ correto.

Resumidamente, da análise das respostas dos alunos nos dois itens, 1a e 2a, constata-se que: 1) em geral, os alunos reconheceram o conceito/definição de DEQ, sendo claro que este é determinado por cinco medidas estatísticas: mínimo, Q_1 , Q_2 (mediana), Q_3 e máximo; 2) muitos alunos revelaram falhas na determinação da mediana e dos quartis (conflitos nos procedimentos), condicionando a obtenção do DEQ correto, sendo que o cálculo do Q_3 , no item 1a, mostrou-se mais complexo devido talvez a, nesse caso, o valor a atribuir a Q_3 não ser um valor da coleção de dados observado no gráfico de barras ($Q_3 = 3,5$); 3) muitos alunos revelaram confusão na forma de apresentar um DEQ ao incorporar o seu suporte numérico ao próprio diagrama (conflito nos conceitos/definições); e 4) perante a necessidade de construir uma escala mais precisa no item 2a, devido a uma maior dispersão nas distâncias que separam as cinco medidas que definem o DEQ, essa precisão foi tendencialmente negligenciada pelos alunos (conflito nos procedimentos). Uma vez que, em ambas as questões, os alunos descrevem,

em geral, corretamente o conjunto ordenado dos dados, infere-se que a forma de representação dos dados não deve constituir um obstáculo à correta construção de um DEQ.

4.2 Interpretação de diagramas de extremos e quartis

Nos itens 1b, 2b e 2c é questionado o comportamento da distribuição dos dados, pretendendo-se confrontar as respostas no item 1b, no qual não é especificado no enunciado o tipo de análise a realizar, com as respostas nos itens 2b e 2c, no qual se solicita que se analise a dispersão e simetria da distribuição dos dados.

Concretamente, o item 1b tem como propósito avaliar a percepção e a capacidade de leitura dos alunos sobre a distribuição dos dados a partir de um DEQ, de forma que o gráfico de barras do enunciado possa, eventualmente, auxiliar na interpretação que resulte dessa leitura. Em particular, esperava-se que os alunos apontassem a assimetria da distribuição (reconhecendo o enviesamento à esquerda), mencionassem a dispersão da distribuição (comentando o domínio de variação dos dados com a indicação de medidas de dispersão, como a amplitude total ou a amplitude interquartil), ou ainda discutissem o posicionamento dos quartis (Q_1 e Q_3 , com referência, por exemplo, a haver pelo menos 75% dos dados não inferiores a Q_1 , ou não superiores a Q_3). Assim, uma resposta no item 1b foi considerada correta, se o aluno tivesse considerado dois daqueles três cenários distintos nas suas conclusões. Com um ou nenhum cenário respondido corretamente, a resposta ao item 1b foi classificada como não correta.

Nos itens 2b e 2c pretendia-se que os alunos comparassem os dois DEQ dados no enunciado, incidindo essa comparação na dispersão (item 2b) e na simetria (item 2c) das distribuições, e apresentando uma narrativa argumentativa com uma linguagem relativamente cuidada e baseada em medidas estatísticas que sustentassem as suas conclusões.

As distribuições das respostas dos alunos segundo os tipos de categoria atribuída (correta, não correta e não resposta), cruzadas pelos itens 1b e 2b, e pelos itens 1b e 2c, encontram-se nas Tabelas 2 e 3.

Tabela 2 – Distribuição (%) do tipo de resposta dos alunos nos itens 1b *versus* 2b

Item 1b	Item 2b			Total
	Correta	Não correta	Não resposta	
Correta (dispersão e simetria corretos)	—	3 (3,2)	2 (2,2)	5 (5,4)
Correta (dispersão e quartis corretos)	—	2 (2,2)	—	2 (2,2)
Correta (simetria e quartis corretos)	—	—	—	—
Não correta (apenas dispersão correto)	—	3 (3,2)	—	3 (3,2)
Não correta (apenas simetria correto)	7 (7,5)	8 (8,6)	—	15 (16,1)

Não correta (nenhum correto)	—	30 (32,3)	6 (6,5)	36 (38,7)
Não resposta	2 (2,2)	2 (2,2)	28 (30,1)	32 (34,4)
Total	9(9,7)	48 (51,6)	36 (38,7)	93 (100)

Fonte: Elaboração dos autores (2018).

Nas respostas ao item 1b, para os tipos de resposta correta e não correta, averiguou-se se os conceitos/definições de dispersão, de simetria ou dos quartis eram abordados, nomeadamente, as respostas com dois cenários corretos (dispersão e simetria; dispersão e quartis; e simetria e quartis), apenas um deles correto (apenas dispersão; apenas simetria) ou nenhum deles correto.

Tabela 3 – Distribuição (%) do tipo de resposta dos alunos nos itens 1b *versus* 2c

Item 1b	Item 2c			Total
	Correta	Não correta	Não resposta	
Correta (dispersão e simetria corretos)	—	—	5 (5,4)	5 (5,4)
Correta (dispersão e quartis corretos)	—	2 (2,2)	—	2 (2,2)
Correta (simetria e quartis corretos)	—	—	—	—
Não correta (apenas dispersão correto)	3 (3,2)	—	—	3 (3,2)
Não correta (apenas simetria correto)	—	7 (7,5)	8 (8,6)	15 (16,1)
Não correta (nenhum correto)	—	34 (36,6)	2(2,2)	36 (38,7)
Não resposta	2 (2,2)	10 (10,8)	20 (21,5)	32 (34,4)
Total	5(5,4)	53 (57,0)	35 (37,6)	93 (100)

Fonte: Elaboração dos autores (2018).

Das Tabelas 2 e 3 constata-se uma elevada percentagem de não respostas (34,3% no item 1b, com 6,5% delas em consequência de não terem construído o diagrama solicitado previamente no item 1a, 38,7% no item 2b e 37,6% no item 2c) e, por outro lado, uma baixa percentagem de respostas corretas (7,6% no item 1b, 9,7% no item 2b e 5,4% no item 2c, sendo nenhuma simultaneamente correta nos pares de itens 1b e 2b ou 1b e 2c). Destas contagens se deduz que existiram muitas dificuldades em responder às perguntas de análise dos DEQ, tendo havido, na questão 2, maior dificuldade em discutir corretamente a simetria do que a dispersão da distribuição dos dados.

No exame das respostas não corretas nos itens 1b (54 alunos), 2b (48 alunos) e 2c (53 alunos), para identificar os objetos primários que determinam as falhas dos alunos na interpretação de DEQ, foram verificados três pontos-chave: percepção intuitiva das noções de dispersão, simetria ou quartis (*conceitos/definições e propriedades*); saber recorrer a medidas estatísticas adequadas (amplitude total ou amplitude interquartil para a dispersão e, por exemplo, as distâncias de cada extremo à mediana para avaliar a simetria) na sua justificação (*argumentos*) e usar uma linguagem com relativo rigor (i.e., com referência a entidades estatísticas corretas) na sua argumentação (*linguagem*).

Relativamente ao item 1b, da Tabela 2 ou 3, constata-se que, dos 61 alunos que responderam a este item, 16,4% (5+2+3 alunos) demonstraram ter percepção do conceito/definição de dispersão, 32,8% (5+0+15 alunos) do conceito/definição de simetria e apenas 8,2% (5 alunos) simultaneamente dos dois conceitos/definições. Esta diferença de percentagens leva a crer que na questão 1, contrariamente ao que se detectou na questão 2, houve menor dificuldade em discutir corretamente a simetria do que a dispersão da distribuição dos dados.

Ainda no item 1b, verificou-se que dos 54 alunos com respostas não corretas, a maioria (36 alunos) apresentaram respostas baseadas na análise do gráfico de barras, sendo muitas delas erradas e não revelando qualquer percepção intuitiva da noção de dispersão ou de simetria de uma distribuição (exemplo na Figura 9). Para além deste, o padrão de resposta mais observado, que denota dificuldade na análise de um DEQ no item 1b, foi a ausência de indicação de medidas estatísticas para descrever a dispersão ou assimetria da distribuição dos dados (27 alunos; um exemplo na Figura 10)

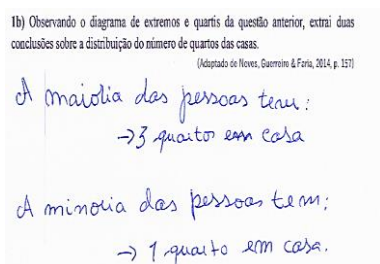


Figura 9 – Resposta do aluno A27 ao item 1b
Fonte: Aluno A27 (2018).

Na resposta apresentada na Figura 9 percebe-se que o aluno se limita a fazer uma análise (errada) do gráfico de barras, destacando os valores da variável com maior frequência absoluta (habitações com três quartos) e com menor frequência absoluta (habitações com um quarto), optando, por fim, pelo valor mais frequente (moda), sem relação com a dispersão ou assimetria.

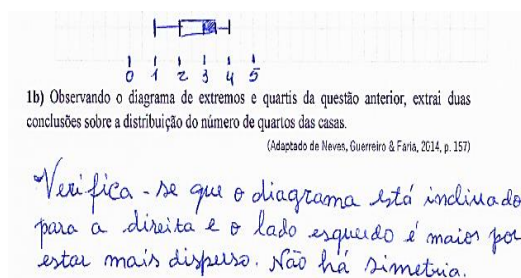


Figura 10 – Resposta do aluno A84 ao item 1b
Fonte: Aluno A84 (2018).

Na Figura 10 encontra-se a resposta do aluno A84 que, embora tenha usado o termo “disperso”, apenas foca a sua análise do DEQ na simetria. Foi pouco rigoroso na afirmação que

redige, não recorrendo a termos estatísticos para descrever as diferenças de amplitudes que intuitivamente descreve para justificar a assimetria da distribuição.

Nas 48 respostas não corretas do item 2b, os padrões mais observados foram: referir, de forma intuitiva, a noção de dispersão, mas não a justificar com uma medida estatística decorrente da observação dos DEQ (14 alunos; um exemplo na Figura 11); considerar que a dispersão é igual nos dois DEQ.

2b) Comparando, quanto à dispersão, as distribuições dos resultados obtidos nos dois primeiros testes de Matemática.

No teste 2 há mais concentração de valores do que no primeiro

Figura 11 – Resposta do aluno A13 ao item 2b
Fonte: Aluno A13 (2018).

Na Figura 11 é apresentada a resposta do aluno A13 que redige uma afirmação correta, assinalando o DEQ com maior dispersão, mas não usa medidas estatísticas para justificar a sua conclusão.

No que concerne ao item 2c (ver Tabela 3), examinaram-se as 53 respostas não corretas, tendo-se observado que os padrões de resposta mais frequentes foram: não usar explicitamente uma medida estatística na sua argumentação de simetria (15 alunos; um exemplo na Figura 12) e trocar a análise da simetria, que era solicitada, pela análise da dispersão da distribuição (13 alunos; um exemplo na Figura 13).

2c) Comparando, quanto à simetria, as distribuições dos resultados obtidos nos dois primeiros testes de Matemática (Adaptado de Costa & Rodrigues, 2014, p.145).

*1º teste → simétrico, pois os valores estão centrados
2º teste → assimétrico, porque os valores estão mais para a esquerda.*

Figura 12 – Resposta do aluno A34 ao item 2c
Fonte: Aluno A34 (2018).

Na Figura 12, a resposta dada pelo aluno A34 revela que ele tem a noção de simetria, mas não recorre a medidas estatísticas para justificar a sua conclusão quanto à assimetria de cada distribuição dada.

2c) Comparando, quanto à simetria, as distribuições dos resultados obtidos nos dois primeiros testes de Matemática (Adaptado de Costa & Rodrigues, 2014, p.145).

O teste 1 é maior do que o teste 2 na variedade de dados e com dispersão também

Figura 13 – Resposta do aluno A52 ao item 2c
Fonte: Aluno A52 (2018).

Na Figura 13, o aluno A52 apresenta uma resposta com pouco rigor na afirmação que redige, focando-se na dispersão e não na simetria como era pedido, para além de não identificar uma medida estatística extraída do DEQ para justificar a sua afirmação.

Comparando as respostas nos itens 1b, 2b e 2c, verifica-se que apenas 5 alunos foram consistentes em diferenciar (intuitivamente) os conceitos/definições de simetria e dispersão nos três itens. Em termos de linguagem usada, embora nem sempre rigorosa, em geral os termos estatísticos foram usados, tendo havido, por vezes, confusão entre simetria e dispersão na argumentação.

Comparando a capacidade de análise intuitiva demonstrada pelos alunos nos conceitos/definições de dispersão e simetria nas questões 1 e 2, conclui-se (ver Tabelas 2 e 3) ter havido uma melhor percepção do conceito/definição de simetria na questão 1 do que na questão 2 (21,5% de respostas corretas no item 1b e 5,4% no item 2c) e desempenhos similares no conceito/definição de dispersão (10,8% de respostas corretas no item 1b e 9,7% no item 2b).

Em síntese, os erros cometidos pelos alunos nos itens 1b, 2b e 2c refletem dificuldades na análise de DEQ, nomeadamente, em associar a percepção intuitiva da forma da distribuição dos dados, resultante da leitura de um DEQ, com a interpretação dessa forma distribucional por meio de medidas estatísticas dos dados que possam suportar essa intuição. Efetivamente, constata-se que: 1) para cada um daqueles itens, mais de um quarto dos alunos opta por não responder; 2) os alunos, quando revelam conhecer intuitivamente a noção de simetria e/ou dispersão de uma distribuição, tendem a não usar medidas estatísticas, obtidas da leitura do DEQ, para justificar a forma distribucional identificada; e 3) as dificuldades detectadas não dependem do rigor da representação gráfica do DEQ, já que elas surgem, quer quando o DEQ é apresentado no próprio enunciado, quer quando este é construído originalmente pelo próprio aluno.

5 Conclusões e reflexões finais

Em termos globais, as ferramentas de análise do EOS (GODINO; BATANERO, 1994; GODINO; BATANERO; FONT, 2007; GODINO et al., 2017) possibilitaram estudar a aprendizagem ocorrida, destacando as ações realizadas e os conhecimentos postos em jogo nas respostas dadas pelos alunos. A análise dessas respostas permitiu identificar os objetos primários e processos usados nas resoluções escritas produzidas pelos alunos às questões 1 e 2 e evidenciar a existência de discrepâncias entre os significados pessoais atingidos pelos alunos e os significados pretendidos. No presente estudo, tais discrepâncias foram mais visíveis nas

alíneas referentes à interpretação de DEQ (itens 1b, 2b e 2c). Todavia, é possível que o fato de os alunos terem sido mais capazes de construir o DEQ terá tido repercussões positivas também na sua interpretação, pois trata-se de processos que interatuam entre si (DINIZ; FERNANDES, 2016).

Na construção do DEQ (itens 1a e 2a) observou-se uma maior facilidade em contexto gráfico do que com dados não organizados, sendo a maior dificuldade dos alunos nos procedimentos, quando da determinação dos quartis (item 1a) e nos conceitos/definições, nomeadamente na apresentação (ou não) do suporte numérico ao desenhar o DEQ (item 2a), com eventual falta de rigor na escala desse suporte.

Estas imprecisões na construção do DEQ, devidas ao suporte numérico, não têm sido salientadas na literatura especializada, ainda que em diversas reproduções de respostas de alunos, tal esteja patente. Para além das aqui reportadas (Figuras 5-8), vejam-se também, por exemplo, em Edwards Özgün-Koca e Barr (2017), um diagrama sem suporte e um suporte com falta de rigor na escala. A falta de cuidado em incorporar corretamente o suporte no DEQ pode advir do fato desta representação não estar num espaço bidimensional no qual habitualmente se representam, com maior cuidado, os gráficos de funções ou não ter sido dada a devida importância pelos próprios professores.

A análise comparativa de distribuições através dos respectivos DEQ (itens 1b, 2b, 2c), no qual se estabelecem semelhanças e diferenças entre padrões das distribuições a nível da dispersão e da simetria, revelou-se muito difícil para a maioria dos alunos, com muitas dificuldades em se expressarem, tal como também verificaram Bakker Biehler e Konold (2004) e Edwards Özgün-Koca e Barr (2017), residindo a maior dificuldade a nível da argumentação.

Os resultados do presente estudo empírico realçam diferenças substanciais de desempenho entre a questão 1 e 2 na análise da simetria da distribuição, o que requer interpelar tal diferença. Para tal, uma análise mais detalhada aos três DEQ em questão mostra que o DEQ no item 1b é mais consistente com a noção de ser “assimétrico” no sentido em que, no domínio de variação total (do valor mínimo ao valor máximo), a distância entre o valor mínimo e a mediana é maior do que a distância entre a mediana e o valor máximo, mantendo-se esta relação de ordem no intervalo interquartil (do primeiro quartil ao terceiro quartil), pois a distância entre o primeiro quartil e a mediana é maior do que a distância entre a mediana e o terceiro quartil (exemplo do DEQ A da Figura 14). Esta consistência não se verifica no DEQ B da Figura 14 (nem nos DEQ dos itens 2b e 2c), no qual, para apenas 50% dos valores mais centrais da distribuição (intervalo interquartil), se verifica que a distância entre o primeiro quartil e a mediana é maior do que a distância entre a mediana e o terceiro quartil, mas no domínio de

variação total, a distância entre o valor mínimo e a mediana e a distância entre a mediana e o valor máximo são visivelmente iguais.

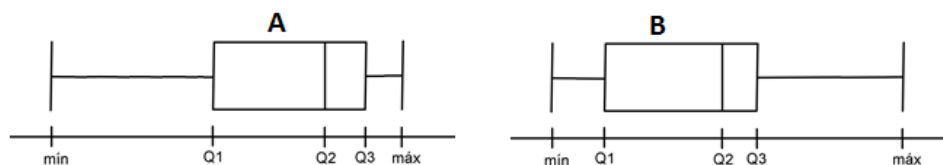


Figura 14 – Diagramas de extremos e quartis. A não simetria é mais intuitiva no DEQ A do que no B
Fonte: Elaboração dos autores (2018).

Mais ainda, verificou-se que no item 1b a noção de simetria foi mais adotada do que a de dispersão. O fato de os dados estarem organizados num gráfico de barras, que é um tipo de gráfico já muito trabalhado, poderá ter permitido ainda aos alunos analisarem o tipo de simetria da distribuição sem recorrerem ao DEQ. Esta situação reforça a dificuldade de interpretar o DEQ, como já foi referido por Bakker Biehler e Konold (2004), diante de um estudo envolvendo alunos do 8.º ano.

Relativamente à análise da dispersão de uma distribuição usando DEQ, foram visíveis dificuldades nas respostas ao item 2b, quando os alunos compararam a dispersão dos dois DEQ associados a duas distribuições com o mesmo número de dados, que consideraram ser iguais ou semelhantes, apesar de não o serem. Morais e Fernandes (2011) apontaram esta como sendo a maior dificuldade sentida por alunos também do 9.º ano.

Por outro lado, verificou-se que, quase sempre, os alunos basearam a sua avaliação da dispersão e da simetria a partir de argumentos intuitivos e raramente apresentaram justificações com base em medidas estatísticas, como sejam a amplitude total, a amplitude interquartil ou outros intervalos relativos a valores típicos do DEQ. Estes argumentos intuitivos, segundo Skemp (1993), surgem da capacidade de combinar e relacionar muitas experiências diferentes, até de classes de experiências.

As maiores dificuldades sentidas pelos alunos na interpretação do DEQ, também observada em outros estudos (EDWARDS; ÖZGÜN-KOCA; BARR, 2017; MORAIS; FERNANDES, 2011), podem ser sucintamente explicadas pelo fato de tal interpretação implicar níveis mais elevados de leitura e interpretação de gráficos (EDWARDS; ÖZGÜN-KOCA; BARR, 2017), como sejam ler entre os dados e ler para além dos dados (CURCIO, 1989).

As baixas percentagens de respostas corretas sobre a dispersão e a simetria de uma distribuição sugerem que não é dada relevância suficiente a esses conceitos/definições no ensino, para permitir ao aluno do ensino básico em Portugal consolidar o seu conhecimento. Será necessário, pois, dar mais tempo para os alunos examinarem diferentes situações,

envolvendo os conceitos/definições de simetria e dispersão, associados àquela representação recém-estudada, como alertam Bakker Biehler e Konold (2004), já que tais conceitos/definições são apenas trabalhados no final do 8.º ano.

A elevada percentagem de alunos (51% dos 191 participantes iniciais) que não responderam a nenhuma questão sobre DEQ, proposto no teste de diagnóstico, demonstra que é necessário refletir sobre o ensino e aprendizagem desse tema. Por outro lado, tendo havido várias respostas do tipo “não conheço”, tal pode indiciar que, seguindo a orientação da maioria dos manuais escolares, o tema foi deixado para o fim do ano letivo, com a agravante de poder ter sido reduzido o tempo previsto para a sua lecionação, em virtude de eventuais atrasados na lecionação do programa.

A interpretação de dados agregados subjacentes à análise de DEQ é também uma outra dificuldade reportada por vários autores (BAKKER; BIEHLER; KONOLD, 2004; EDWARDS; ÖZGÜN-KOCA; BARR, 2017). Efetivamente, o DEQ é um tipo de representação aparentemente simples, mas que não é simples de analisar. Importa articular atividades nas quais os alunos verifiquem que diferentes coleções de dados podem originar iguais DEQ, como é ilustrado na Figura 3, e verificar que nem sempre é simples, por exemplo, tipificar a simetria da distribuição dos dados através de um DEQ. Estarão os professores alertados para este detalhe?

Assim, emerge deste estudo a necessidade de articular atividades envolvendo diferentes cenários. Uma vez que grande parte do conhecimento dos alunos é apreendido diretamente daquilo que os rodeia, o que contribui para transformar conceitos abstratos em menos abstratos, uma sugestão relevante para a aprendizagem e o ensino dos DEQ consistirá em, numa fase inicial, começar por explorar as ideias intuitivas dos alunos acerca da dispersão e simetria em cenários significativos, acompanhado da determinação de estatísticas que validem ou questionem as suas intuições e, prosseguindo, na fase seguinte, com a diversificação de cenários mostrando diferentes coleções de dados que originam o mesmo DEQ.

Por fim, os resultados obtidos neste estudo sugerem também a realização de outros estudos complementares que possam contribuir para melhorar a aprendizagem dos variados conceitos/definições associados à construção e interpretação dos DEQ.

Agradecimentos

Trabalho subsidiado por fundos portugueses através do CIDMA (Centro de Investigação e Desenvolvimento em Matemática e Aplicações) da Universidade de Aveiro e FCT (Fundação para a Ciência e a Tecnologia), dentro do projeto UID/MAT/04106/2019, e também pelo CIEd

– Centro de Investigação em Educação, UID/CED/01661/, Instituto de Educação, Universidade do Minho, através de fundos nacionais da FCT/MCTES-PT; EDU2016-74848-P (AEI, FEDER).

Referências

ARTEAGA, P.; BATANERO, C.; CAÑADAS, G.; CONTRERAS, J. M. Las tablas y gráficos estadísticos como objetos culturales. **Números**, La Laguna, v. 76, p. 55-67, 2011.

BAKKER, A.; BIEHLER, R.; KONOLD, C. Should young students learn about box plots? In: BURRILL, G.; CAMDEN, M. (Ed.). **International Association for Statistical Education 2004 Roundtable**: Curricular development in statistics education. Voorburg: International Statistical Institute, 2004. p. 163-173.

BOAVENTURA, M. G.; FERNANDES, J. A. Dificuldades de alunos do 12.º ano nas medidas de tendência central: O contributo dos manuais escolares. In: FERNANDES, J. A.; SOUSA, M. V.; RIBEIRO, S. A. (Org.). **Ensino e aprendizagem de probabilidades e estatística** – Actas do I Encontro de Probabilidades e Estatística na Escola. Braga: Centro de Investigação em Educação da Universidade do Minho, 2004. p. 103-126.

COSTA, B.; RODRIGUES, E. **Novo Espaço**: Matemática 8.º ano/Parte 2. Porto: Porto Editora, 2014.

CURCIO, F. R. **Developing graph comprehension**: elementary and middle school activities. Reston: NCTM, 1989.

DINIZ, L. N.; FERNANDES, J. A. Interações entre construção e interpretação de gráficos estatísticos em projetos de modelagem matemática com uso de tecnologias de informação e comunicação. **VIDYA**, Santa Maria, v. 36, n. 2, p. 457-475, 2016.

EDWARDS, T. G.; ÖZGÜN-KOCA, A.; BARR, J. Interpretations of boxplots: Helping middle school students to think outside the box. **Journal of Statistics Education**, Alexandria, v. 25, n. 1, p. 21-28, 2017.

ESPINEL, M. C.; GONZÁLEZ, M. T.; BRUNO, A.; PINTO, J. Las gráficas estadísticas. In: SERRANO, Luis R. (Ed.). **Tendencias actuales de la investigación en educación estocástica**. Málaga: Gráficas San Pancrancio, 2009. p. 57-74.

FERNANDES, J. A.; BARROS, P. M. Dificuldades de futuros professores do 1.º e 2.º ciclos em estocástica. In: CONGRESSO IBERO-AMERICANO DE EDUCAÇÃO MATEMÁTICA (CIBEM), 5, 2005, Porto. **Anais...** Porto: Faculdade de Ciências da Universidade do Porto, 2005.

FREITAS, A.; CRUZ, J. P.; SILVA, N. Mediana de dados não agrupados: a questão de ser pelo menos 50%. **Educação e Matemática**, Lisboa, n. 143, p. 18-21, 2017.

FREITAS, A.; FIGUEIREDO, T.; SILVA, N.; MIRANDA, M. C. Dificuldades na aprendizagem da Mediana e Quartis por alunos do 8.º ano de escolaridade: estudo comparativo Fórmula versus Gráfico. **Indagatio Didactica**, Aveiro, v. 10, n. 2, p. 109-132, 2018.

FRIEL, S.; CURCIO, F.; BRIGHT, G. Making sense of graphs: critical factors influencing comprehension and institutional implications. **Journal for Research in Mathematics Education**, Reston, v. 32, n. 2, p. 124-158, 2001.

GAL, I. Adult's statistical literacy. Meanings, components, responsibilities. **International Statistical Review**, Voorburg, v. 70, n. 1, p. 1-25, 2002.

GALL, M.; GALL, P.; BORG, W. **Educational research**: An introduction. Boston: Allyn and Bacon, 2003.

GODINO, J. D. Indicadores de la idoneidad didáctica de procesos de enseñanza y aprendizaje de las matemáticas. In: CONFERÊNCIA INTERAMERICANA DE EDUCAÇÃO MATEMÁTICA, 13., 2011, Recife. **Anais...** Recife: Universidade Federal de Pernambuco, 2011.

GODINO, J. D.; BATANERO, C. Significado institucional y personal de los objetos matemáticos. **Recherches en Didactique des Mathématiques**, Grenoble, v. 14, n. 3, p. 325-355, 1994.

GODINO, J. D.; BATANERO, C.; FONT, V. The onto-semiotic approach to research in mathematics education. **ZDM - The International Journal on Mathematics Education**, Verlag, v. 39, n. 1-2, p. 127-135, 2007.

GODINO, J. D.; GIACOMONE, B.; BATANERO, C.; FONT, V. Enfoque ontossemiótico de los conocimientos y competencias del profesor de matemáticas. **Bolema**, Rio Claro, v. 31, n. 57, p. 90-113, 2017.

LOPES, C. E. O ensino de probabilidade e estatística na escola básica nas dimensões do currículo e da prática pedagógica. In: SIMPÓSIO IBEROAMERICANO DE ENSEÑANZA MATEMÁTICA, 16., 2004, Castellón. **Anais...** Castellón: Universitat Jaume I, 2004.

MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA. **Programa e Metas Curriculares de Matemática do ensino básico**. Lisboa: Autor, 2013.

NEVES, M. A. F. **Matemática 8**. Porto: Porto Editora, 2014.

MORAIS, P. C.; FERNANDES, J. A. Realização de duas tarefas sobre construção, leitura e interpretação de gráficos estatísticos por alunos do 9º ano. In: SEMINÁRIO DE INVESTIGAÇÃO EM EDUCAÇÃO MATEMÁTICA (XXII SIEM), 22., 2011, Lisboa. **Anais...** Lisboa: Associação de Professores de Matemática, 2011.

OLIVEIRA, H.; HENRIQUES, A. Promover o raciocínio estatístico no ensino básico recorrendo à tecnologia. **Boletim da SPE**, Lisboa, Outono, p. 23-31, 2014.

PAIS, S.; SARAIVA, M. J. O Significado das representações da função afim para discentes do 8.º ano de Escolaridade. **Quadrante**, Lisboa, v. 20, n. 2, p. 17-55, 2011.

PONTE, J.; FONSECA, H. Orientações curriculares para o ensino da estatística: Análise comparativa de três países. **Quadrante**, Lisboa, v. 10, n. 1, p. 93-115, 2001.

RUMSEY, D. J. Statistical literacy as a goal for introductory statistics courses. **Journal of Statistics Education**, Alexandria, v. 10, n. 3, 2002.

SERRANO, L.; ORTIZ, J.; RODRIGUEZ, J. La simulación como recurso didáctico en la enseñanza de la probabilidad. Las gráficas estadísticas. In: SERRANO, L. R. (Ed.). **Tendencias actuales de la investigación en educación estocástica**. Málaga: Gráficas San Pancraccio, 2009. p. 157-176.

SKEMP, R. **Psicología del aprendizaje de las matemáticas**. 2. ed. Madrid: Ediciones Morata, 1993.

TUKEY, J. W. **Exploratory data analysis**. Reading: Addison-Wesley, 1977.

Submetido em 25 de Agosto de 2018.
Aprovado em 07 de Agosto de 2019.